Australian Government
Department of Defence
Defence Science and
Technology Organisation

Automated Detection and Classification in High-resolution Sonar Imagery for Autonomous Underwater Vehicle Operations

Philip Chapple

Maritime Operations Division
Defence Science and Technology Organisation

DSTO-GD-0537

ABSTRACT

Autonomous Underwater Vehicles (AUVs) are increasingly being used by military forces to acquire high-resolution sonar imagery, in order to detect mines and other objects of interest on the seabed. Automatic detection and classification techniques are being developed for several reasons: to provide reliable and consistent detection of objects on the seabed; to free human analysts from time-consuming and tedious detection tasks; and to enable autonomous in-field decision-making based on observations of mines and other objects. This document reviews progress in the development of automated detection and classification techniques for side-looking sonars mounted on AUVs. Whilst the techniques have not yet reached maturity, considerable progress has been made in both unsupervised and supervised (trained) algorithms for feature detection and classification. In some cases, the performance and reliability of automated detection systems exceed those of human operators.

RELEASE LIMITATION

Approved for public release

Published by

Maritime Operations Division
DSTO Defence Science and Technology Organisation
PO Box 1500
Edinburgh South Australia 5111 Australia

Telephone: (08) 8259 5555
Fax: (08) 8259 6567

AR-014-199
December 2008

Automated Detection and Classification in High-resolution Sonar Imagery for Autonomous Underwater Vehicle Operations

Executive Summary

Autonomous Underwater Vehicles (AUVs) are increasingly being employed for mine reconnaissance, mine hunting and hydrographic survey operations. Side-looking sonar systems can generate high-resolution seabed imagery, indicating the presence of mines and other bottom objects. Whilst human analysts may be tasked to examine the data, this approach is resource-intensive and potentially unreliable, as analysts become tired or inconsistent in their performance and are often distracted by other tasks.

This document reviews the development of techniques for automated detection and classification of objects on the seabed from this imagery. These techniques have been developed to provide more reliable and consistent detection of significant objects, in order to free operators from these time-consuming and tedious detection tasks. Automatic detection and classification also enable real-time sonar processing to take place onboard suitably equipped AUVs, allowing for autonomous decision-making based on current observations.

Techniques for computer-aided detection/classification (CAD/CAC) in sidescan sonar imagery have been under development since the early 1990s, principally in North America and Europe. The most successful techniques rely on the presence of a coupled acoustic highlight and shadow associated with an object sitting proud of the seabed. The challenge has been to develop algorithms that can detect and classify mine-like objects reliably, with very few false alarms. The performance of these algorithms depends on the sonar system, the background clutter and other prevailing environmental conditions, which can significantly influence the observability of target objects in sonar imagery.

Two broad classes of detection/classification algorithm are in use: supervised algorithms, requiring training data with target objects in known locations, and unsupervised algorithms. Well-designed supervised algorithms can be expected to have superior performance for particular environments when trained with appropriate data. The main limitation in applying these algorithms is that suitable training data sets are not always available or easy to acquire. The training data must be extensive and obtained under similar sonar and environmental conditions to those in the data for which object detection is required, but in the training data the actual distribution of mine-like objects must be known. Unsupervised algorithms are designed to work under a range of conditions, in the absence of training data. They are therefore simpler to implement operationally, without the requirement for additional surveys to obtain suitable training data.

Fusing the results of several different algorithms can dramatically improve the performance of CAD/CAC systems over the performance using any one of these algorithms on its own. Different methods of fusing the results have been tested and enhanced detection probabilities demonstrated, with acceptably low false alarm rates. In order to achieve significant gains, it is necessary for these algorithms to perform fundamentally differently from one another. Using this approach, CAD/CAC performances exceeding human performances have been observed.

Synthetic aperture sonar (SAS) has the operational advantage of allowing for high-resolution surveys of the seabed with an increased detection range, enabling AUVs with these sonars to survey the seabed more rapidly. CAD/CAC techniques developed for sidescan sonar have also been applied to SAS imagery. While the shadows in SAS imagery are less distinctive and there are some other differences from conventional sidescan, processing techniques are being developed to allow objects in SAS imagery to be readily detected by automated processing.

For post-processing of seabed imagery, it remains to be seen whether CAD/CAC systems will be trusted to take the place of human analysts. For this to happen, the success of these systems must be demonstrated for a range of operational and environmental conditions. It is envisaged that, once these systems are trusted, they will be routinely employed to highlight areas of images that warrant close inspection by a human analyst, obviating the requirement for the analyst to scan through all the data. This procedure will greatly increase the speed and efficiency of mine countermeasures operations and other operations requiring seabed feature detection.

When CAD/CAC systems are incorporated into real-time processing systems on board AUVs, the vehicles will be able to make autonomous decisions based on detection of seabed features. An AUV could be programmed to respond to the presence of a mine-like object in one of several ways: by returning to the location of the object for a closer inspection with higher-resolution sensors; by tasking another vehicle to examine the object in more detail; or by transmitting information about the object back to a control platform. This technology is likely to provide a significant enhancement to the effectiveness of naval mine countermeasures and underwater survey operations.

Contents

1. INTRODUCTION 1

2. OPERATIONAL ADVANTAGES OF AUTOMATED IMAGE PROCESSING..... 2

2.1 Automation as a decision aid 2

2.2 Automation and unmanned maritime vehicles 4

2.3 Other uses for automated image processing 5

3. FEATURES OF SIDE-LOOKING SONAR IMAGERY 5

3.1 Scanning sensors 5

3.2 Sidescan sonars 5

3.2.1 Identification of contacts in sidescan imagery 8

3.2.2 Geometrical and natural factors impacting on CAD/CAC............. 10

3.2.3 Equipment design and CAD/CAC.. 12

4. IMAGE ENHANCEMENT 13

5. COMPUTER-AIDED DETECTION AND CLASSIFICATION 15

5.1 Supervised methods 16

5.1.1 US Navy sponsored research... 17

5.1.2 Canadian research 18

5.2 Unsupervised methods 19

5.3 Recent work 20

6. FUSION OF ALGORITHMS 21

7. CAD/CAC IN SYNTHETIC APERTURE SONAR IMAGERY................................ 23

7.1 Introduction to synthetic aperture sonar 23

7.2 SAS imagery 24

8. CONCLUSION 26

9. ACKNOWLEDGEMENT.. 28

10. REFERENCES 28

1. Introduction

Making sense of imagery is something that comes naturally to humans, but it remains a challenge to provide a similar capability to computers and robotic systems. Nevertheless, computational image processing has progressed rapidly in the last twenty years, enabled by developments in image processing techniques and software and by rapid advances in sensors and computer performance.

The emergence of robotic systems has been a key driver for developments in computational image processing. Unmanned vehicles, particularly autonomous vehicles, are particularly benefited by advances in image processing, as they are thereby enabled to make decisions about their environments in order to navigate and perform their tasks. Image analysis potentially enables a mobile robot or autonomous vehicle to respond to the presence of objects, plan complicated navigational paths and avoid collisions.

Image processing is an enormous field of research with many potential applications to unmanned vehicle systems. This report considers image processing techniques that are primarily relevant to unmanned maritime vehicle systems tasked with naval mine hunting and route surveillance operations; ultimately, such vehicles require capabilities for autonomous detection and characterisation of mine-sized objects on the seabed and in the water column.

At present, high-resolution side-looking sonar systems, such as sidescan sonar (SSS) and synthetic aperture sonar (SAS), are the tools of choice for imaging the seabed to detect mines and mine-like features. Sonars of this type and various high-resolution optical and laser imaging systems also feature as the main tools for further classification and identification of detected objects. Large data volumes are an inherent consequence of the use of high-resolution imaging systems. More often than not, the communications links available on remotely operated or autonomous systems lack sufficient bandwidth to transmit such data off-board in real or close-to-real time. Consequently, it is often not possible for a human analyst to have enough information to make a timely decision about the best course of action. Communications bandwidth is a particular constraint on the operation of Autonomous Underwater Vehicles (AUVs); there is insufficient bandwidth in underwater acoustical or electromagnetic communications channels to support rapid transmission of sonar data, so imagery is typically stored on board the vehicle, to be downloaded and processed after its mission is complete.

The capability to process high-resolution imagery on board an unmanned vehicle is highly desirable, to give the vehicle an autonomous decision-making capability and also to augment the capability of humans involved in image analysis. But while the vehicle navigation and guidance technologies have reached the point where unmanned marine surveys have become routine, automated image analysis techniques are not mature. Many approaches to image analysis are available and they vary widely in their speed, efficacy, resource requirements, accuracy and robustness. Hence, there is a need to examine the available techniques, and to employ and develop techniques applicable to Australian Unmanned Maritime System (UMS) operations.

In 2000, Perry [1] reviewed the applications of image processing to mine warfare sonar operations. The current document updates that work and concentrates more specifically on high-resolution sidescan and synthetic aperture sonars of the kind used in unmanned maritime vehicles. The processing of forward-looking sonar imagery is not considered here, because forward-looking imaging sonars are not currently available in most autonomous maritime systems. This is not to imply that such sonars are not worthy of study if and when they become available. It should be noted that somewhat different techniques are appropriate to the processing of data and imagery from such sonars.

Section 2 describes in more detail the military operational advantages of automated sonar image processing for UMS operations. In Section 3, features of side-looking sonar imagery are described, as it is these features that determine the kind of processing that is suitable for computer-aided detection (CAD) and classification (CAC).[1] Section 4 describes different approaches to pre-detection image enhancement. The development of CAD/CAC processing techniques is surveyed in Section 5, and advantages of fusing different algorithms are discussed in the following section. Some differences apply in the CAD/CAC processing of SAS imagery, as described in Section 7. Finally, overall conclusions and implications for future research by DSTO and the Australian Defence Organisation are presented in the final section.

2. Operational advantages of automated image processing

The need to maintain maritime freedom of manoeuvre implies a requirement for a capability to survey shipping lanes, ports and harbours and to detect and identify sea mines and other objects of significance which might threaten safety of navigation. Currently, this capability is provided through a variety of manned assets and clearance diving teams. However, for reasons of safety, economy and efficiency, unmanned vehicles are increasingly being used as complementary or alternative tools for such tasks.

Automated image processing has the potential to make major contributions to the task of detecting and characterising small objects, particularly for mine reconnaissance and mine hunting operations.

2.1 Automation as a decision aid

In the near term, automation has the potential to reduce the burden on human analysts engaged in the post-mission analysis of large volumes of sonar and other sensor data recorded by high-resolution sensors. The importance of this capability will increase as the resolution of the data increases. Put simply, analysis of seabed imagery is a tedious, time-consuming task requiring considerable attention on the part of the operator. Computer-aided detection/classification (CAD/CAC) of objects in sonar imagery can free operators to concentrate on complex tasks, such as mine identification and disposal, rather than more

[1] The terms 'ATD' (automatic target detection) and 'ATR' (automatic target recognition) are also in use; 'ATR' is commonly used as an alternate to 'CAD/CAC'.

routine image inspection and analysis. Automation potentially enables faster, more consistent processing of the data, eliminating problems of variable performance caused by operator distraction or fatigue.

Partial automation, whereby operators are alerted by the CAD/CAC system to the presence of mine-like objects (MLOs) and other significant features within the data, can also be valuable as a means of reducing the data that the operators must visually inspect to relatively limited areas of concern. This process is more rapid and reliable than relying on personnel to go through all the unprocessed imagery, *provided* the probability of detection of significant features is acceptably high and the probability of false alarms is acceptably low. A useful rule-of-thumb is that, for an automated detection system to be trusted, the expectation of detecting a genuine target must be at least ten times the expectation of encountering a false alarm [2].

Pitfalls in this process have been described in detail by Kessel [2-4]. In many cases, where CAD/CAC systems are intended to assist an operator in detecting targets, these systems come to be regarded more as a burden than an aid. This situation arises when CAD/CAC systems and human operators analyse the same data, but come to different conclusions about the presence of valid targets. This 'second opinion' places additional cognitive burdens of deliberation and ambiguity on the decision-makers, which they find unhelpful. A more satisfactory approach is to have a CAD/CAC system that performs a simple task with high reliability, so that the job of going through all the data is left to the system alone. Such a system can be designed to detect regions of interest to be passed to a human operator for investigation. Imagery from only these regions is passed to the operator, thereby avoiding confusion or conflict between the judgements of the CAD/CAC system and operators in other parts of the data.

It is difficult to create a CAD/CAC system that is trusted sufficiently by human operators to ensure its regular operational use. When such a system is being tested operationally, comparisons are often made between the detections of the CAD/CAC system and those of a human operator. Kessel [2] has identified and quantified problems that can arise in this supervised automation process, caused by:

(i) human operators performing better than the machine at the detection process and rendering the CAD/CAC process unreliable; and

(ii) operators themselves performing unreliably at difficult detection processes, and hence being unable to recognise high-quality performance of a CAD/CAC system.

Both of these scenarios can lead to the CAD/CAC system being rejected. A possible solution to this conundrum is to have independent, objective means of quantifying the performance of a CAD/CAC system, such as assessments of performance in detecting known targets — not relying solely on comparisons with the performances of human operators.

2.2 Automation and unmanned maritime vehicles

A major attraction of unmanned maritime vehicles in mine warfare applications is that all types of vehicle diminish the risks inherent to personnel and high-value platforms working in a minefield. In terms of automated image processing, further advantages accrue from the nature of autonomous surface and underwater platforms:

1. Image quality. AUVs and actively-stabilised surface-towed sensor platforms provide exceptionally stable, uniform platforms for high-resolution sensors. In addition, both types of platform can be operated in 'terrain-following' mode, whereby their altitude above the seabed remains approximately constant and image resolution and contrast remain at near optimal levels throughout the mission.
2. Access to the underwater environment. Unmanned Maritime Vehicles (UMVs) are typically much smaller than manned platforms with equivalent sensing capability. They are thus considerably more manoeuvrable. In the case of AUVs, manoeuvre in constricted areas and close to facilities is practical, as is close-range survey of deep waters.
3. Capability for clandestine operations. UMVs, particularly AUVs, equipped with automated image processing capabilities, provide some degree of clandestine mine detection and characterisation capability.

Automated image processing has a particular role to play in AUV operations, as it can enable intelligent onboard decision-making based on acquired imagery. For example, the detection of a mine-like object or another seabed feature of interest could trigger an AUV to return to the site of the object for a more thorough inspection with a higher resolution sensor. Also, if desired, the AUV could surface to transmit target information back to base. Similarly, the ability of an AUV to detect shoals, coastlines and underwater hazards could enable it to modify its trajectory and hence travel safely in relatively unknown areas.

In the longer term, reliable real-time processing of imagery from mine-hunting platforms has the potential to reduce the total human effort required to clear an area of mines, through increased automation of the entire process. Unmanned vehicles with real-time processing can potentially work together with other manned and unmanned platforms to cover mine reconnaissance, hunting and clearance tasks. Real-time processors already exist for some AUVs and tasking of AUVs by other AUVs has already been demonstrated, but, by common agreement, the reliability of the process is not yet sufficient for it to be operationally useful. Better image processing technology is required.

2.3 Other uses for automated image processing

The primary image processing capabilities being assessed in this study are detection and characterisation of target objects, principally sea mines. Automated image processing techniques and related methods can also yield related capabilities that can contribute significantly to the overall capability of the system; for example:

- A capability to infer sediment characteristics such as roughness, acoustic reflectivity/scattering strength and mechanical shear strength is useful as a means of identifying those areas where object detection and characterisation are likely to be difficult; for example, soft sediments where mines may bury.
- A capability to identify and map features of the seabed and in marine structures can be useful when it is necessary to find small objects in cluttered or constricted areas such as wharves and coral reefs. 'Change detection', involving the comparison of recent and historical data, can assist in the detection of newly placed hazards or threats, even in cluttered areas.
- A capability to estimate the bathymetry and topography of the underwater environment can be useful for navigation, and as an input to the survey planning process.

3. Features of side-looking sonar imagery

3.1 Scanning sensors

Sidescan and synthetic aperture sonars are two of a variety of side-looking, scanning sensors, including multibeam echosounders and laser scanners, which can be used to explore the seabed and the water volume in detail. Rather than imaging a two-dimensional area with every data cycle in the way a camera would, scanning sensors look sideways and downwards, sensing the environment in a vertical plane. This information is projected onto a line drawn along the seabed. The data from a single scan line is a record of reflected intensity as a function of range or, in some cases, as a function of angle. The motion of the platform then provides a second dimension, perpendicular to the first. If the platform is moving in a straight line at uniform speed, the scan-lines are parallel and build up a 'raster chart' of the seabed. If the scanner looks on both sides of the platform, a two-sided image is acquired, doubling the rate of coverage.

3.2 Sidescan sonars

Figure 1 shows an idealised view of the operation of sidescan sonar. The sonar moves along a straight 'track' at constant speed and altitude; that is, constant height above the seabed. Transducers on either side of the sonar send out narrow fans of energy localised around planes perpendicular to the direction of motion; that is, 'across-track'. Port and starboard sides of the imagery thus originate from separate sensors. Raw sidescan imagery corresponds to acoustic echo intensity versus time of flight (echo return time since the 'ping' was emitted), or equivalently, 'slant range'. The horizontal range can be deduced from the slant range by assuming that the seabed is flat and level, to either side. A track of individual sonar scan-lines is referred to as a 'swath'. The region directly under the sonar is referred to as the 'nadir'.

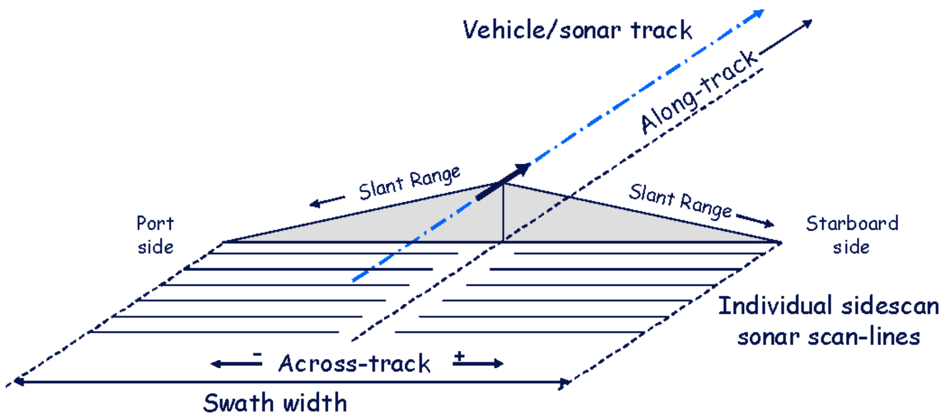

Figure 1: Operation of a sidescan sonar

Sidescan sonar has no resolution in elevation angle; that is, echoes from a given range produce almost the same response in the sonar regardless of the elevation (vertical) angle from which they originated. This is illustrated in Figure 2, which shows an end-view of the acoustic energy emitted from a sidescan sonar and the echoes that it generates.

Echoes originating directly from the seabed constitute the 'signal'. Echoes from the sea surface and arriving at the sonar via multiple bounces from the seabed or sea surface constitute unwanted 'reverberation'. The regions underneath the sonar – the 'nadir' – and above the sonar – the 'zenith' – correspond to points of exceptionally high reflectivity from the seabed and sea surface, respectively. As horizontal range on the seabed is estimated from slant range, the nadir is also the point at which the range resolution of the sonar is lowest and the distortion of the imagery is greatest. In addition, many sidescan sonars preferentially 'ensonify' angles close to the horizontal. Near-vertical angles may be unevenly ensonified or not ensonified, producing a stripe or intensity variation corresponding to the steepest angles.

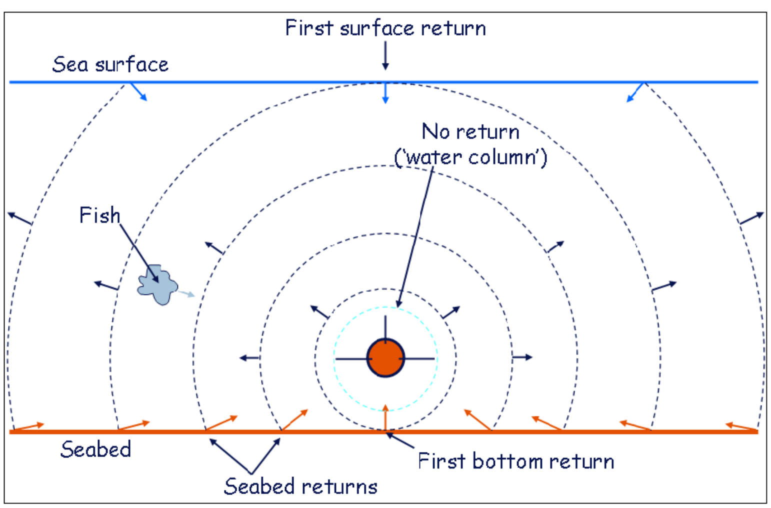

Figure 2: End-view of a sidescan sonar, showing echoes originating on the seabed, at the sea surface and in the water column

Figure 3 shows a typical segment of 'waterfall' imagery from a high-resolution sidescan sonar. The sonar is moving up the page. Slant range, or equivalently time of flight, increases from the centre line to the left and right for the port and starboard channels, respectively. The bright band in the centre of the image corresponds to the emission of the ping. Other range-dependent features are common to both sides of the imagery. The dark strip from 0 to 3 m is the period of low return when the sound is travelling through the 'water column'. The 'first bottom return' at 3 m is followed by some light and dark ripples extending to approximately 7 m due to non-uniformities in the beam pattern of the transducers, when operating at 3 m altitude. The 'sweet spot' of this sonar extends from approximately 10 m to 30 m, the edge of the image. A faint line at approximately 16 m corresponds to the first surface return – little other evidence of surface reverberation is visible in this image, although it may be more significant if the image were collected in choppy conditions. The remaining features in the imagery correspond to the texture of the seabed, which consists of alternating bands of exposed, rippled sand and thick, linear mats of 'line-weed'.

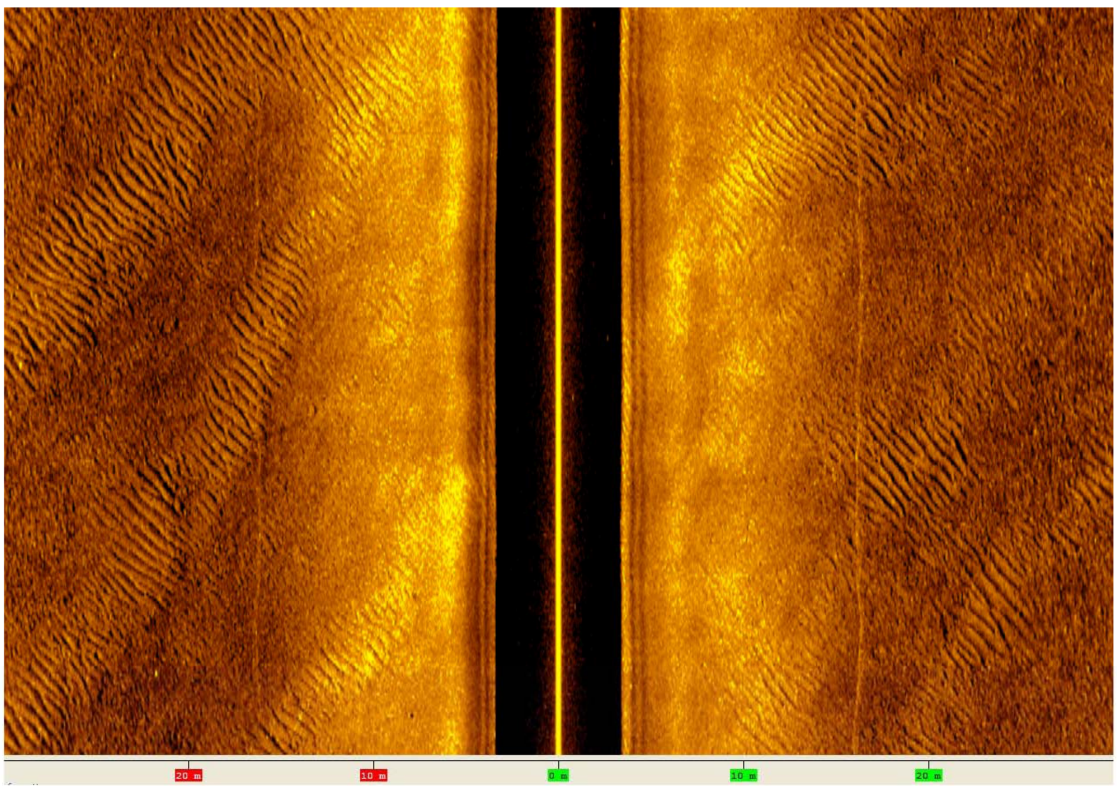

Figure 3: Imagery from a 900 kHz Marine Sonar sidescan sonar installed on a REMUS 100 AUV. Image intensity corresponds to sonar echo intensity.

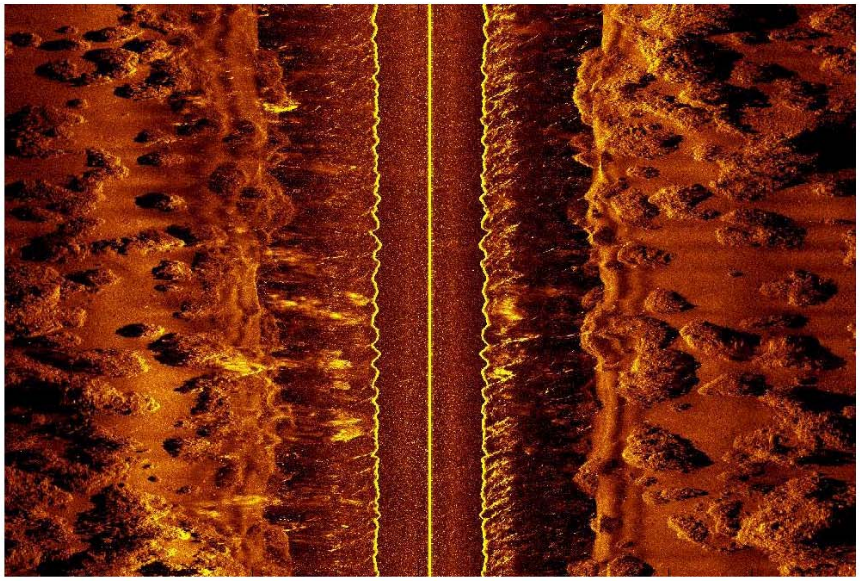

Figure 4: Imagery from the same sonar, exhibiting seabed clutter and surface reverberations

A rather different image from the same sonar is shown in Figure 4. In this case, the seabed is highly cluttered, containing features corresponding to coral outcrops. The AUV was closer to the sea surface than to the bottom, so the strong linear feature from the first surface (zenith) return is closer to the centre line than is the first bottom return. The region between the first surface return and first bottom return shows strong surface reverberations, dependent on sea-state.

3.2.1 Identification of contacts in sidescan imagery

Some sidescan sonars are able to provide imagery with pixel resolutions of a few decimetres or better, suitable for detecting mines and other objects on the seabed. Objects protruding above the seabed are typically considerably more reflective than the surrounding sediment, so a bottom object is often associated with a high-intensity 'highlight' in the sonar imagery. In this sense, a sonar image is similar to a sector-imaging radar image. However, an important additional characteristic of sidescan sonar imagery is that objects that protrude above the seabed block the passage of sound to the sediment behind them, thereby casting distinctive 'shadows' – areas of echo intensity considerably lower than the background level arriving from the seabed. The length of the shadow depends on the vertical extent of the object, relative to the sea floor. Figure 5 shows a high-resolution image of a small boat equipped with an outboard motor. Although the highlights in the image give a good deal of information about the nature of the wreck, only its shadows give an indication of its three-dimensional shape.

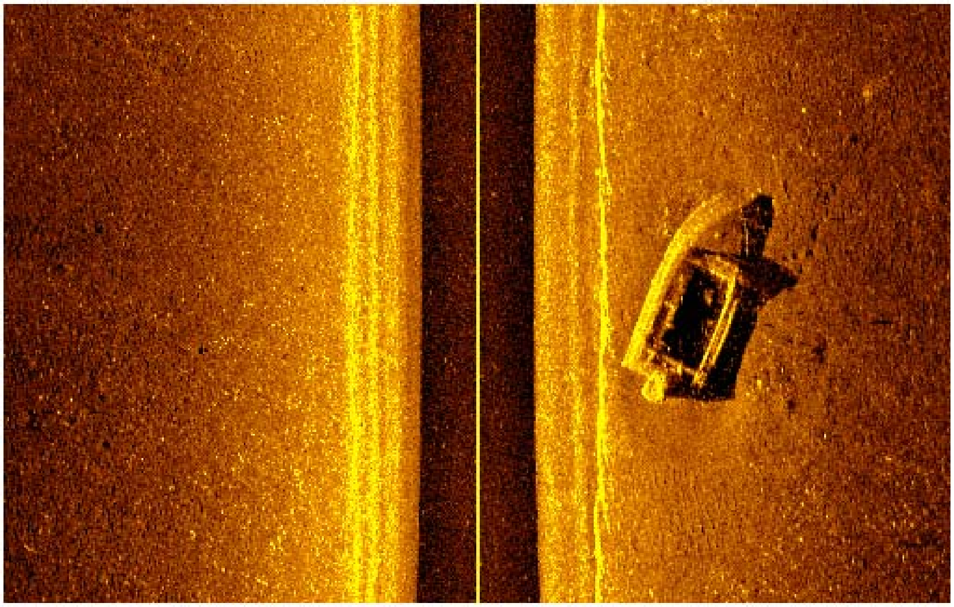

Figure 5: Imagery of a small boat wreck from a 1800 kHz sidescan sonar installed on a REMUS 100 AUV, showing highlights, shadows and decimetre-level resolution.

For a human analyst or a CAD/CAC process, the presence of a highlight in a certain size range, together with an adjacent shadow, reliably signals the presence of a mine-like object (MLO). Figure 6 shows two examples of highlight-shadow contact detection, one recorded with sub-decimetre resolution and one at half the resolution and in poorer conditions.

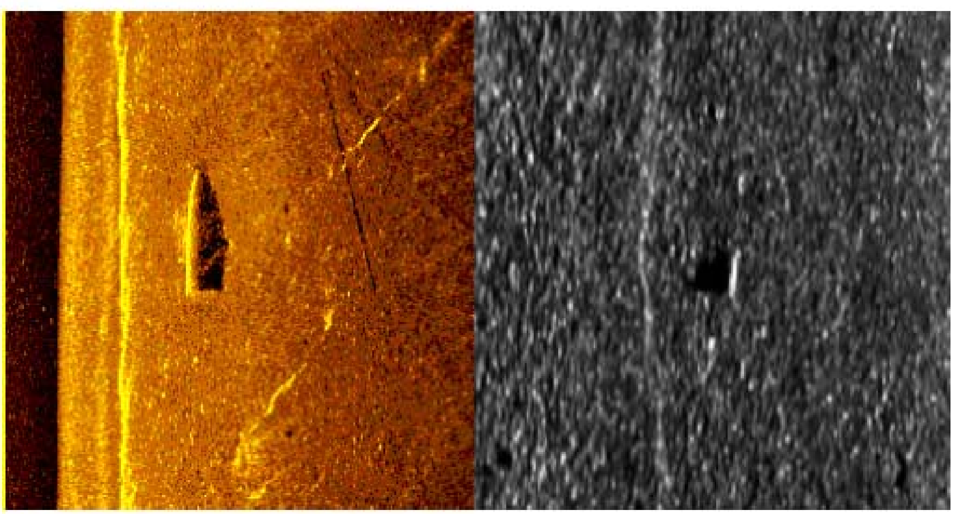

Figure 6: (Left) Imagery of a mine-shape from a 1800 kHz sidescan sonar installed on a REMUS 100 AUV, showing highlight and adjacent shadow. Note the presence of reverberation from surface waves as striations in the imagery. (Right) Imagery of a similar mine-shape from a 900 kHz sidescan sonar installed on the same vehicle.

Other types of detection are also possible. Depending on the relative orientation of the sonar to the object, the strength of a highlight may vary considerably and may indeed fall below the detection threshold of the sonar and therefore be invisible. Despite this, if the geometry of the sonar relative to the object is favourable, a shadow may be present even if a highlight is not, because the passage of sound is blocked by the object. Consequently, it is not uncommon for the shadow associated with an object to be the only indication of its presence. At the other extreme, some sonar-object geometries are not favourable to the formation of shadows. If the horizontal distance from the sonar to the object is not at least two or three times the sonar's altitude above the seabed, the shadow may not extend far enough from the object for it to be distinguishable. Alternatively, if the water depth is much less than the range of the object from the sonar, then 'shadow infill' may occur, whereby signals arriving at the sonar via intermediate surface bounces are approximately as intense as the direct signal from the seabed and cause the same response in the sonar; in effect, the sonar sees reflections of the seabed in the sea surface and vice-versa. In this case, the contrast of the shadow to the background intensity may be reduced or eliminated, so that highlights are the only option for object detection.

Standard operating procedures for sidescan sonar object detection surveys are designed to maximise the benefits of shadow detection. Sidescan sonars are typically flown at an altitude of one tenth of the (per side) range setting so that the region where shadow lengths are small is only a small fraction of the total extent of the imagery. In addition, an infill-line survey pattern is a standard technique that is adopted when mine detection is an important component of the survey mission. Primary survey lines are separated by a distance equal to twice the range setting of the sonar. Secondary 'infill' survey lines are then placed parallel to the primary lines and offset by one-half of the range setting, thereby ensuring that the nadir region of each primary line falls within the sweet spot of each secondary line, and vice-versa. Maximum ranges are also sometimes restricted to avoid shadow infill.

3.2.2 Geometrical and natural factors impacting on CAD/CAC

Because of the simplicity of the scanning process, sidescan sonar imagery is prone to numerous unwanted geometrical and natural artefacts that may interfere with the CAD/CAC process.

1. Turns. Waterfall imagery is geometrically consistent with the seabed only when the sonar is travelling in a straight line. Imagery recorded when the vehicle is turning is strongly distorted and must be identified and discarded.
2. Biological clutter, primarily fish. The swim bladders of fish are efficient scatterers of sound at most frequencies used for sidescan sonar imaging and the bodies of fish can block the higher frequencies. Consequently, dense schools of fish can give rise to contacts with compact highlights and well-defined shadows.
3. Shadow-inducing terrain, primarily sand ripples. When sand ripples are oriented within 45° of the vehicle track orientation, they may give rise to alternating highlights and shadows that have many of the characteristics of contacts associated with mine-like objects. Ridges, reefs and holes may have a similar appearance.
4. Surface reverberation. Reflections from the zenith and from surface wave facets oriented towards the sonar may give rise to strong, compact highlights, although they

are unlikely to be associated with shadows. Whitecaps and bubble trains can be particularly strong scatterers of sound.

5. Burial. Most high-resolution sidescan sonar frequencies have little or no significant ability to penetrate marine sediments. Consequently, objects that are partially buried may lack shadows and vary considerably in appearance from objects that are proud (lying on the seabed) and objects that are fully buried become completely undetectable.

6. Clutter. Numerous natural and man-made objects such as rocks and packing crates may appear mine-like when ensonified from particular angles. All such objects are valid mine-like contacts in the absence of further information; such information may be supplied by ensonification from different angles or at higher resolution, or some other form of inspection may be necessary.

7. Seabed variability. The seabed itself is subject to wide variations in composition, acoustic reflectivity and texture, all of which affect sidescan sonar imagery and the appearance of contacts with respect to the seabed.

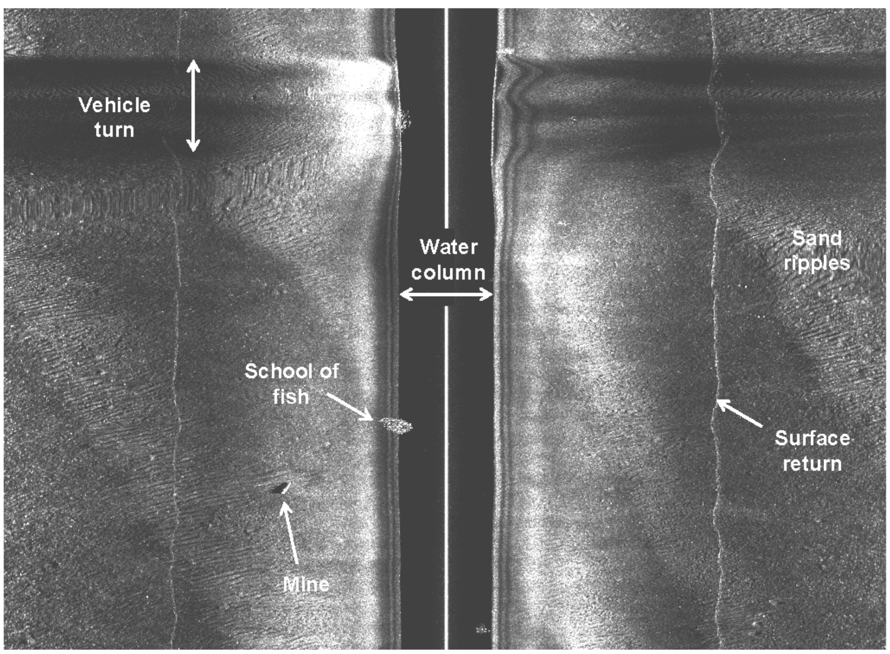

Figure 7: Sidescan sonar imagery from a 900 kHz Marine Sonic sonar installed on a REMUS 100 AUV, showing a mine-shape and various distracting features

Figure 7 shows the effects of turns, fish and surface reverberation on sidescan sonar imagery containing a mine-like contact. Figure 8 shows a mine-like object located amid sand-ripples, which have similar acoustic characteristics.

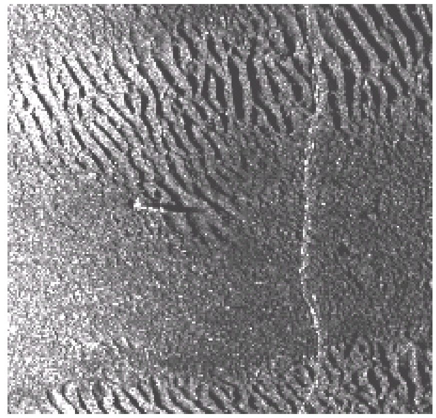

Figure 8: Sidescan sonar imagery from a 900 kHz Marine Sonic sonar installed on a REMUS 100 AUV, showing a mine shape located among sand ripples, and a surface reverberation effect.

3.2.3 Equipment design and CAD/CAC

It cannot be emphasised too strongly that starting with a good data set is vital to achieving success (high probability of detection and classification P_{dc} and low probability of false alarm P_{fa}) in CAD/CAC. It is difficult for any detection and classification process, whether human or computer-based, to work well with noisy, reverberation-dominated, distorted or poorly-resolved imagery. Investments in stable sonar platforms and high-resolution, high-contrast sonars are therefore critical to the success of the mission as a whole. Starting with good data gives subsequent processing a much greater chance of success.

Resolution. Other things being equal, increasing resolution usually makes detection and classification of contacts easier [5]. Various strategies have been attempted in order to increase the resolution of sidescan sonars in azimuth (along-track).[2] Two of these strategies are: increasing the operating frequency; and introducing long, multi-element transducer arrays with relatively sophisticated beamformers. The first approach has achieved decimetre and sub-decimetre resolution as the frequency has approached and exceeded 1 MHz. There is a trade-off, because of the increasing acoustical attenuation of seawater as the frequency increases. Consequently, only limited ranges are attainable at higher frequencies — say, 20,000 times the wavelength, or 30 m at 1 MHz. The second approach has achieved 1 to 2 decimetre resolution with frequencies of order 500 kHz, at perhaps double the range but with much greater cost and complexity. At present, the first approach is predominant, but the second is also practicable, albeit at a higher cost and with a physically larger sonar head. Considerable advances in range and resolution are expected as synthetic aperture sonar processing becomes a mature field, allowing lower frequencies to be used to achieve both ranges extending beyond 100 m and sub-decimetre resolution.

[2] It is easier to achieve high resolution in the across-track direction (by high-speed temporal processing) than along-track (involving angular or spatial processing).

Contrast. Optimal function of CAD/CAC algorithms relies on there being sufficient 'dynamic range' or contrast in the imagery to accommodate the highlights due to strong acoustic returns, the various mean intensity levels of the seabed and the shadows due to occlusion of acoustic energy. Successive generations of sidescan sonar have incorporated progressively less noisy amplifiers, particularly at higher frequencies, and digitisers with wider dynamic range. Nevertheless, the imagery shown in preceding figures was all collected with Marine Sonic sonars recording data with only 6-bit digitisation, whereas some higher-end sonars employ 8, 12, 16 and even 24-bit [6] digitisation. By careful attention to automatic gain control (AGC) and time-varying gain (TVG) to preserve useful contrast across the image, compact 6-bit digitisers, such as in the Marine Sonic sidescan images shown in the figures, can remain effective. Nevertheless, processing can be improved with higher fidelity data and the use of digital, rather than analogue, filtering techniques [6]. These improvements come at the expense of greater cost and complexity of the sonar systems and much greater volumes of data to be stored and processed.

Platform stability. As already noted, AUVs and actively-stabilised towbodies are optimal platforms for the collection of sidescan sonar imagery, in terms of their ability to maintain straight, uniform motion at a set altitude above the seabed.

Reverberation. Image 'clutter' corresponding to uninteresting objects on the seabed is unavoidable, but sonar systems can be designed and equipment operated to minimise the impact of surface and volume reverberation on imagery. Careful attention to the shape of the main lobe of the sonar transmit/receive beams and to reduction of sidelobes can reduce the unwanted reverberation and maximise the effectiveness of the sonar. By operating an AUV well below the surface and ideally at times of low sea state, surface reverberation effects can be reduced. Multipath effects including the infilling of acoustical shadows can be reduced by ensuring that data are collected at a suitable range, using the operating procedures described in Section 3.2.1.

4. Image enhancement

Various processes referred to as 'image enhancement' may be applied to imagery as a pre-processing step prior to application of CAD/CAC algorithms. Such processes are intended to make the tasks of detection and classification easier by removing obvious artefacts and outliers from the imagery.

The simplest form of image enhancement corrects for the variation of image intensity with range from the sonar, which may otherwise impact on the optimal thresholds for detection of targets above background clutter. This variation is partly corrected by TVG and AGC, as mentioned in the previous section, but such corrections, when performed by real-time sonar processors, are frequently less than perfect. The dependence of seabed reverberation with range depends on the frequency of the sonar, its altitude above the seabed and the seabed type. A rough and rocky seabed will reflect a significant fraction of the incident acoustic wave back towards the sonar, whether the wave arrives at nadir or grazing incidence, whereas a flat sandy seabed will reflect back strongly at nadir incidence, but weakly at grazing incidence.

A common normalisation technique that results in an image intensity that is, on average, constant with range, operates by dividing the image intensity by an average intensity for that range. Care is necessary if this technique is to be worthwhile. As can be seen, for example, in Figure 3, the water-column region is quite different from the remainder of the image. Simple range-dependent normalisation will not be ideal if the sonar altitude (and hence water-column width) changes within the image. To deal with this scenario, it can be useful to form a 'slant-range-corrected' image, which is resampled so that the horizontal position on the image corresponds to the horizontal distance from the nadir (assuming the bottom is flat in the across-track direction). Image normalisation can then be carried out on the resulting imagery. Image normalisation is not always required — it depends on the CAD algorithm. In the author's early work [7] it was not carried out, because the detection process involved dividing each image into subimages for processing, based on the statistics of those subimages — performing the normalisation to some extent as part of the detection process.

A further image enhancement that may be useful is to normalise not only the mean but also the variance of pixel intensities, for each range, prior to processing. Furthermore, in previous work on CAD in land imagery [8], pixel intensity values were scaled to convert non-Gaussian distributions to Gaussian ones (histogram distortion). This was done because the CAD algorithms used in that work were optimal for a Gaussian distribution of pixel intensities.

Speckle noise reduction is a form of image enhancement that is sometimes used in sidescan image processing, to remove scintillation caused by coherence effects in the scattered sound.[3] The aim is to remove noise spikes without impairing the capability to detect targets. Johnson [9] investigated median filtering versus morphological filtering (nonlinear filters involving erosion, dilation, opening and closing operations) to reduce speckle noise. Median filters are often used, but are more computationally costly than morphological filters, which can achieve a comparable level of performance.

Several more refined statistical techniques have been used to generate enhanced imagery that minimises high-spatial-frequency noise in sidescan sonar imagery, while preserving features of interest. The Total Variation Minimisation technique [10-12] minimises a functional[4] that has one term favouring image smoothness (low intensity gradient) and another term favouring faithful replication of the recorded imagery, including features such as highlights and shadows. This approach has been demonstrated to improve CAD/CAC performance significantly. Huynh *et al.* [13] used the wavelet transform effectively to reduce high-spatial-frequency clutter in sidescan sonar imagery, improving detection performance and reducing false alarm rates. For optimal mine detection performance, the scale of the wavelets should be matched to the size of expected mines.

---

[3] Speckle noise arises due to constructive interference between different scattering points in the footprint of the sonar beam, like the speckle that is visible at a point illuminated by a laser pointer.

[4] A functional is a mapping from a vector space or a space of functions to (usually) real numbers. In the TVM technique, the functional is minimised to obtain the function most suitable for reducing the high-spatial frequency noise in the imagery.

An adaptive clutter suppression linear filtering technique has been investigated by Aridgides *et al.* [14-15] as a precursor to CAD/CAC processing. This technique requires training data — images containing background only and images containing targets. A small window is moved across the image, and local values of the covariance matrix are calculated and used to evaluate filter coefficients to suppress clutter without suppressing targets. This technique is optimal in a least-squares sense, for Gaussian-distributed clutter. It is more complex and computationally demanding than the methods of image enhancement described earlier. It does, however, effectively lower the probability of false alarms (P_{fa}) in CAD/CAC processing, without adversely affecting the detection probability (P_d).

An important overall systems approach is to collect data such that the requirement for pre-detection image enhancement is minimised. For complex and cluttered marine environments, pre-detection image enhancement will always be advantageous, provided it does not impose too great a computational burden. This computational burden is a consideration if rapid, real-time processing is desired.

5. Computer-aided detection and classification

The object of this report is an examination of the relative merits of certain algorithms for computer-aided detection and classification. This is not straightforward, for a number of reasons:

- The terms 'detection' and 'classification' are not well-defined. The act of detection necessarily involves an element of classification — sufficiently mine-like or not to warrant further investigation[5] — that must then be resolved by a further 'classification' step. Judgements about what constitutes a mine-like object may thus influence the statistical estimates of the probability of detection/classification (P_{dc}) of mine-like objects and the probability of false alarm (P_{fa}).
- Performance is environmentally dependent. Certain CAD/CAC algorithms work well in particular conditions, such as for detection of mines lying on flat sandy seabeds, where the signal-to-noise ratio (SNR) is high, but other algorithms may out-perform them for high clutter, low SNR situations. It is therefore difficult to come up with a single approach that is universally applicable.
- It is also difficult to make an objective comparison of algorithms that have been run on different sidescan sonar data sets. Quantitative comparisons are valid only when different algorithms are applied to the same data sets, with the same definitions of mine-like objects (MLOs) and false alarms.
- The standard way of quantifying the performance of a CAC/CAC system is by means of the Receiver Operating Characteristic (ROC) curve, in which the probability of detection/classification P_{dc} is plotted against the probability of false alarm P_{fa} [16], but often authors do not report the performance of their algorithms in these terms. It is therefore difficult to compare performances of different algorithms quantitatively.

[5] An operational approach to labelling a mine detection that is adopted by some naval officers is to ask 'Would I drive my boat over it?'

Noting these considerations, the following analysis of algorithms is descriptive, rather than being numerically based.

Techniques for computer-aided detection and classification (CAD/CAC) of mine-like objects (MLOs) in high resolution sonar imagery have been the subject of concentrated effort in North America and Europe since the early 1990s [11-47 and references therein]. The general approach is two-pass: firstly, detect target objects in the imagery with a high probability of detection and a high probability of false alarm; and secondly, classify detected targets into MLO and non-MLO categories in order to achieve a much lower total probability of false alarm.

Because of the distinctive shadows that are cast by a sidescan sonar, the most successful CAD/CAC algorithms in use all rely on the correlation of the intensity highlights from bottom objects with the shadows of these objects. In fact, as pointed out in Section 3.2.1, the shadows generally appear more consistent than the variable highlights from objects of interest. Therefore, shadows have a primary role in the detection and classification of man-made objects on the seabed.

The various detection/classification techniques that have been developed can be broadly divided into two groups: unsupervised algorithms and supervised learning algorithms.

5.1 Supervised methods

Supervised detection/classification algorithms are 'trained'; that is, they are optimised so as to locate a set of previously identified mine-like objects within a training data set. The performance of supervised algorithms is highly dependent on the nature of the training data set. Ideally, such a data set should contain numerous combinations of backgrounds and MLOs viewed from different ranges and aspect angles. However, it is not necessarily true that the training operation should always employ the entire data set. An algorithm that is trained with one sort of seabed background, or one particular sonar, may perform poorly when applied to data containing a different kind of background. Likewise, an algorithm trained for one type of sonar or sonar setting may perform poorly when used with data collected from another. The point at which a training data set becomes 'sufficiently large' is difficult to define, but there must be sufficient variety in the training data to ensure that correct classification performance depends not on anomalies in individual images in the training data, nor on peculiarities of the particular training data set. The training data set should be representative of all possible appearances and orientations of mines and backgrounds in the 'test data' – that is, the data in which detection of MLOs is ultimately required. Improvements in detection and classification have been observed [17] by using different training data sets for different scenarios (different image resolutions and SNR values), rather than using an aggregate training data set covering all possible scenarios.

Because it is difficult to acquire suitable training data in sufficient quantities, some researchers have generated their own data synthetically, and have inserted mines at random locations, with random orientations [17, 32, 36-41]. The mines must be inserted with shadows that are realistic and take into account the acoustic angle of incidence and the topography of the seabed. With a sufficiently large data set, receiver operating characteristic (ROC) curves can be generated showing the probability of detection/ classification (P_{dc}) as a function of the probability of false alarm (P_{fa}). The training data set is presumably sufficiently large when the ROC curve does not vary significantly when further data are added to the training data set, or when particular members of the training data set are removed. A hybrid technique that is sometimes used is to insert mine-like contacts artificially into real sidescan sonar imagery. This approach obviates the requirement to synthesise realistic seabed imagery matching the test data, while allowing for much more imagery containing mines than can readily be obtained by field measurements.

The best results in detection and classification for a given sonar data set can potentially be obtained by fusion of the results from several different algorithms, as will be discussed in more detail in Section 6.

5.1.1 US Navy sponsored research

Pioneering research on CAD/CAC processing of sidescan sonar imagery has been undertaken since the early 1990s by Dobeck and others [13-15, 17-31] at the US Naval Surface Warfare Center Coastal Systems Station (CSS), in collaboration with colleagues at Lockheed Martin, Raytheon and Colorado State University. Dobeck *et al.* [19] initially enhanced the image by background normalisation, followed by convolution of the image with nonlinear matched filters as a first-pass detector for MLOs. Filter masks were chosen according to the expected mine type and the background statistics, taking into account the highlight-shadow pairings associated with real mines. Targets were detected by scanning a target-sized window over the normalised matched-filtered image, and counting pixels that exceeded a certain threshold.

Following this first-pass detection, for each of the candidate targets, up to 45 feature statistics were calculated, pertaining to the size and shape of the highlight and shadow. Optimisation procedures were used as part of the training process, to determine the best combinations of features to use to build multidimensional feature vectors for classifying MLOs. Note that it is not always necessary or desirable to use all the available features; using a smaller number of mutually independent features is better than using a large number of features that are interrelated.

Dobeck *et al.* used two different classifiers to decide whether initially detected targets were mines: a *K*-nearest neighbour neural network (KNN) and an optimal discriminatory filter classifier (ODFC). These techniques are both supervised, and hence require training data in order to establish classification criteria. The KNN technique involves a two-layer neural network, which classifies features according to the proximity of the feature vectors to 'feature vector centres' for each classification class. The ODFC is a classifier based on linear discrimination theory, using linear filters based on the characteristics of the mines and the background clutter. The KNN and ODFC classification results were then combined by Boolean AND to yield the final classification result for each target. This fusion of classification results gave rise to better classification performance (or lower false alarm rates) than either technique could achieve individually, because of the fundamental differences between the two techniques. Advantages of fusing different algorithms will be discussed further in the Section 6.

An adaptive filter technique was employed at Lockheed Martin by Aridgides *et al.* [14-15], for detection and classification of MLOs in sidescan imagery, based on a Bayesian classifier known as the log-likelihood ratio test (LLRT) [21]. Again, this CAD/CAC technique requires training of the algorithms using images containing targets and backgrounds. An average target signature (normalised shape) was estimated from training images containing targets and a background clutter covariance matrix was calculated from images containing backgrounds. A two-dimensional linear filter was then formed, optimal in a least squares sense, to preserve features resembling the average target peak signature while repressing background clutter, and this filter was applied to the test data.

The initial work by Lockheed Martin was improved and extended to include pre-processing, adaptive clutter filtering, image normalization and detection, extraction of feature vectors, orthogonalisation of these vectors and optimised classification using LLRT [22-23]. The final result was a correct mine classification and false alarm rate performance that was better than that obtained by an expert human sonar operator.

Also in the 1990s, Raytheon developed CAD/CAC techniques for processing imagery from the AN/AQS-20 helicopter-towed minehunting system and the REMUS AUV [22-26]. These techniques were based on median filtering to reduce speckle, followed by image segmentation, feature extraction, classification and identification of contacts.

The different CAD/CAC approaches of CSS, Lockheed Martin and Raytheon are summarised in Table 1. Several schemes of fusing these different algorithms have been attempted, as described in Section 6.

Table 1: Comparison of three US CAD/CAC algorithms (from [22])

CSS	Lockheed Martin	Raytheon
– Image normalisation – Nonlinear matched filter detector – Feature extraction – Optimal feature selection – KNN attractor-based neural net – Optimal discrimination filter classifier	– Adaptive clutter filter detector – Feature extraction – Feature orthogonalisation transform – Optimal subset feature selection – Log-likelihood ratio test classifier	– Multi-stage median filtering – Image normalisation – Highlight/shadow segmentation – Invariant shape-based features – Multi-level scoring-based classification

5.1.2 Canadian research

Fawcett [32] developed a supervised technique whereby small image sections containing mines are used as the feature vectors for target detection and classification. Principle component analysis (PCA) was used to identify the most significant image features to characterise the different images and variations between them, reducing the size of the feature vectors. This is a commonly used technique for facial recognition. Discriminant analysis was then used to recognize differences between the feature vectors pertaining to different object classes (manta-like, cylindrical and rock). Linear and quadratic classification techniques were trialled successfully on synthetic images. In [33], this approach was applied to real trials data and compared with the use of feature vectors derived from highlight/shadow segmentation and analysis. It was found that both approaches worked well, but that the best results came from a combination of highlight/shadow analysis and PCA.

5.2 Unsupervised methods

Training of CAD/CAC processes to recognise mines on the seabed has advantages, in that the methods are optimised to perform well under the training conditions, but there are also disadvantages. The main issue is that while the process might work well for one kind of sonar and seabed, it is not guaranteed to perform well when the range, resolution or seabed appearance is quite different. Because of this limitation and the lack of sufficient quantities of suitable data for training the algorithms, some researchers have chosen to use untrained (unsupervised) methods. Note that training data is often difficult to obtain — it requires images of the seabed that are similar to and representative of those in which target detection is required, with known targets available for training the algorithms. For mine detection operations in new and untested areas, the acquisition of suitable training data may not be practical.

Unsupervised algorithms are generic; that is, they must work for a broad range of input data, and they are not optimised for any particular set of training data. They might not work as well as an algorithm that is trained on the same kind of data for which detection is required, but they have the advantage of broad applicability without the requirement for suitable training data.

A Markov Random Field (MRF) model of the seabed background was developed by Mignotte *et al.* [34-35]. This model is able to describe seabeds that include sand waves and other clutter or structure, and allows for segmentation of the image into different texture regions. In this work, computationally intensive methods such as simulated annealing and a genetic algorithm were tested for their ability to detect objects on the seabed. The genetic algorithm gave more favourable results.

Reed *et al.* [37-38], at Heriot Watt University and SeeByte Ltd, have used an MRF model to segment sidescan sonar images into three different regions: highlights (including returns from bottom objects), shadows of objects and general background. The segmentation is direction-oriented; it takes account of the fact that the shadow of an object protruding from the seabed will fall on the long-range side of the object. While determining the optimal MRF parameters is computationally intensive, approximations can be made to speed up the process. Post-segmentation processing is used to select highlights of a mine-like size which are paired with neighbouring shadows.

For extraction of object features and classification, a cooperating statistical snakes model is used to identify the boundaries of objects and shadows. This method is an extension of a standard technique for segmenting images to isolate objects, by enclosing them in snake-like boundaries. In the cooperating statistical snakes model, the highlight and the shadow are enclosed in this way with two boundaries that are constrained to be mutually consistent. In this way, realistic feature boundaries are able to be drawn for both the highlight and the shadow, enabling classification of mine-like objects in the presence of sand waves, which can disturb the mine and shadow boundaries calculated using other algorithms. In further work [39-40], Dempster-Shafer theory (an extension of probability theory based on 'belief functions') was used as an aid in the classification process. This approach helps in the classification of objects which may have been viewed multiple times from different aspects.

5.3 Recent work

More recently, the Heriot-Watt group has used a supervised algorithm for detection and classification, based on features calculated using central filters [41]. Training of the algorithm enabled features to be classified as mine or non-mine. A major emphasis of this work was the use of 'augmented reality' images to provide the large number of images including target objects that are required for training. In this approach, targets were synthetically placed in real sidescan sonar imagery, at random positions and orientations. A seafloor model was constructed from the sidescan imagery [42], and this information was used to calculate likely appearances of the targets with their shadows. This approach was found to be effective, enabling the trained algorithms to detect real MLOs in trials data.

Science Applications International Corporation (SAIC — Newport, RI, USA) has employed CAD/CAC to automate the processing of large quantities of sidescan data, collected commercially for the National Oceanic and Atmospheric Administration (NOAA) [44]. Constant False Alarm Rate (CFAR) detection is used, with a split window to detect a highlight followed by a shadow. Sand waves are mitigated by Fourier transforming the images to place them in the wave-number domain, in which periodic sand waves give rise to peaks, which are then removed using a median filter. A neural network scheme was trained to classify the detected objects into mine/non-mine categories.

Chapple [7] has used a straightforward, unsupervised approach to detection of mines in high-quality imagery obtained from DSTO's REMUS 100 AUV. This technique makes use of the fact that, in many images containing mine-like objects, some of the brightest pixels in that part of the image correspond to returns from the mines, while some of the darkest pixels in a local area correspond to shadows. Images are divided into small sections, in which the local intensity histograms are calculated to determine highlight regions (pixels occupying the top few percentiles of the histogram) and shadow regions (pixels occupying the bottom few percentiles). Highlight and shadow regions within specified size limits are then identified, and highlight/shadow pairs satisfying certain geometrical relationships are regarded as detections. When applied to high-quality imagery, this approach yielded few false alarms. Further development and testing are required to compare the performance of this simple technique with statistical analytical approaches implemented in commercial software [45].

6. Fusion of algorithms

While individual CAD/CAC algorithms have their strengths and weaknesses, it is often possible for a combination of algorithms to perform significantly better than any one algorithm in isolation [28]. That is, for a given false alarm rate, there will be a higher probability of detection.

In order to gain from the use of multiple algorithms, it is necessary that the different algorithms are, to a significant degree, statistically independent of one another. Detection/ classification algorithms a and b are said to be statistically independent if the joint probability $P_{dc}(a,b)$ of detecting and correctly classifying an MLO in both algorithms is equal to product $P_{dc}(a)P_{dc}(b)$ of the individual detection/classification probabilities. This condition has been observed to be reasonably accurate in practice [28], when applied to the results of algorithms that operate quite differently. Similarly, differently operating algorithms often give rise to different false alarms, so that the probability that both algorithms will generate the same false alarm is relatively small.

Studies by Aridgides *et al.* [22-23] considered the fusion of the three detection/ classification algorithms from Lockheed Martin, Raytheon and the US Naval Surface Warfare Center, Coastal Systems Station, described in Section 5.1.1. Various methods of fusion were investigated for detection/classification probabilities and the numbers of false alarms. These methods included 'logic-based fusion' methods, in which the three sets of results were combined using various combinations of the results combined using Boolean AND and OR operators. Another successful method was the '2-out-of-3' method, a particular instance of m-out-of-n fusion ($m <= n$). This means that if there are n algorithms, and a target is detected and classified as an MLO by at least m of these algorithms, then the target is included in the overall result.

Aridgides *et al.* found that significant improvements over these methods can be obtained by employing the log-likelihood ratio test (LLRT) algorithm in fusing the results of different detection schemes. In this approach, detection confidence vectors are formed and feature vector orthogonalisation is performed, so that optimal decision rules can be formulated. LLRT-based fusion exhibited a threefold reduction in the false alarm rate over the 2-out-of-3 method, and a 4:1 improvement over logic-based fusion [23], as shown in Figure 9. Further recent improvements in fusion techniques are described in [30].

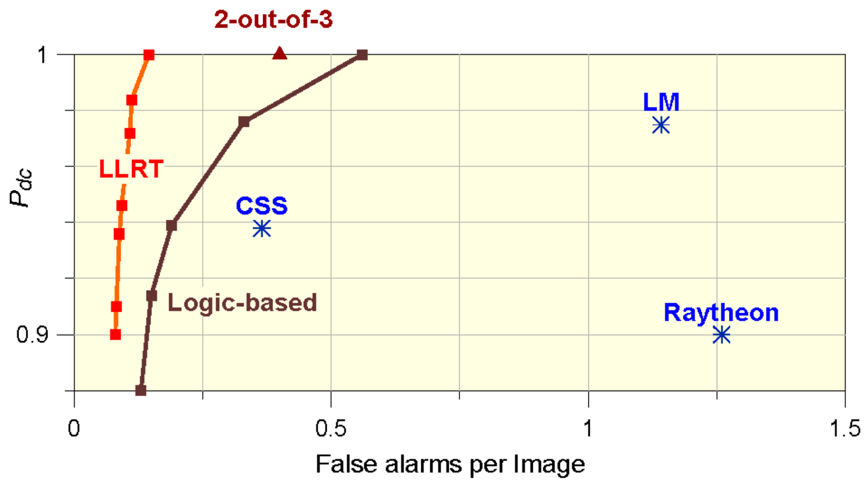

Figure 9: Results of fusion of CAD/CAC algorithms, for one set of input data (from [23])

In score-based algorithm fusion, the k-th detection algorithm assigns a score s_k between 0 and 1 to any object that it detects. If the score is greater than a certain threshold value, then a contact is regarded as having been detected; otherwise it remains undetected. Fusion of algorithms can be performed by a number of processes, such as comparing the total of scores $\{s_k\}$ for an object with a threshold value, or using some other linear combination of the scores to calculate a weighted sum.

A recent study by Dobeck quantified the gains that are available in a score-based fusion technique [31]. Dobeck found that the probabilities P_d and P_f of detection and false alarm in his scenario are approximately given by

$$P_{d\text{-}fusion} = P_{d\text{-}min}\,; \qquad\qquad P_{f\text{-}fusion} = 2^{-(n-1)}P_{f\text{-}min}\,;$$

where n is the number of fused algorithms, and $P_{d\text{-}min}$ and $P_{f\text{-}min}$ are the minimum P_d and P_f values of all the algorithms. Thus, by fusing four or five algorithms, one could hope to reduce the false alarm rate to one eighth or one sixteenth of the best P_f value, without any loss in detection performance over the worst-performing algorithm. In this scenario, one can afford to run the individual algorithms with higher P_d and higher P_f than would normally be tolerated, in the knowledge that the fusion process will bring the false alarm rate down.

7. CAD/CAC in synthetic aperture sonar imagery

7.1 Introduction to synthetic aperture sonar

In the past few years, the resolution and range of sidescan sonar have arguably approached the limits of what is technically and operationally feasible. One-to-two decimetre and sub-decimetre resolutions have been achieved at the cost of long, multi-element transducer arrays, or by moving to frequencies around 1 MHz or higher, but at these frequencies range is severely limited by acoustic absorption in the water [46]. For example, DSTO's REMUS 100 AUV is fitted with a simple Marine Sonic sidescan sonar that has a 900 kHz channel with a resolution of 20 cm and a practical maximum range of 30 to 40 m, and a 1.8 MHz channel with a resolution of 5 to 10 cm and a maximum range of 10 to 15 m. A 2.4 MHz sonar from the same manufacturer that reputedly achieves 1 cm resolution has a maximum range of only around 6 m [47]. The larger and more complex L-3 Klein 5500 sidescan, which operates at 455 kHz and includes a 12-element, 1.2 m long transducer array, achieves 20 cm resolution out to a range of about 75 m. Range and range resolution enhancements have also been achieved by moving to wide-band, pulse-compression signal processing. Ultimately, however, the scale and difficulty of the mine detection problem suggests the desirability of sonars capable of achieving sub-decimetre resolutions extending over swath widths much in excess of the few metres to perhaps few tens of metres that are currently feasible. The array lengths necessary to achieve this with conventional sidescan sonar become inconvenient and unwieldy, especially when the sonar is required to fit on the hull of an AUV.

For more than two decades, synthetic aperture sonar (SAS) has been investigated as a potential solution to the limitations of conventional sidescan sonar, and development has reached the point where a few models with potentially desirable characteristics have become commercially available [48-50]. In synthetic aperture processing, echo returns from a series of sonar pings are combined so that there is effectively an aperture (transducer array length) that is much longer than the transducer array *element* length l. While the maximum cross-track range is the same as for a conventional sidescan sonar operating at the same frequency, it is theoretically possible to maintain the along-track resolution independent of cross-range by aperture synthesis. In principle, for a single transducer element and an unlimited effective aperture, the along-track resolution can be maintained at $l/2$, as for synthetic aperture radar (SAR).

In practice, for the sonar to function as from a synthetic aperture, the position of the sonar at each ping must be known with great accuracy and the difficulty of maintaining sufficient positional accuracy increases as the aperture length increases. The resolution that can be achieved is therefore 1.5 to 2 times coarser than the theoretical value [48] and tends to degrade slowly with range. In addition, factors such as electronic and ambient noise become more important as range increases, and multipath reverberation also increases, to the point that the maximum effective range of the sonar may be dictated by reverberation in shallow water.

A further difference between synthetic and real-aperture sonar is the function of multiple-element receive arrays. In a real aperture sonar, the length of the receive array determines the resolution achievable by the sonar and the maximum speed of advance. In synthetised aperture sonar, the total length of the aperture determines only the maximum speed of advance. In essence, the sonar cannot travel more than half the length of the receive array per ping interval, limiting the attainable range for a given array size and platform speed.

The engineering problems associated with SAS processing have proven to be more difficult to solve than those associated with SAR, which is now widely used. Nevertheless, the capabilities of SAS devices now being marketed are impressive: for example, the 100 kHz HISAS 1030 [48] sonar developed by FFI, the research arm of the Norwegian Department of Defence, in conjunction with Kongsberg Maritime, is claimed to be able to achieve better than 5 cm resolution both along-track and cross-range for ranges up to 200 m at 4 knots. It is also interferometric, so the resulting imagery is associated with accurate bathymetry [51]. A further advantage of the HISAS sonar is that Kongsberg Maritime claim to have succeeded in making their HUGIN 1000 AUV sufficiently stable to accommodate the HISAS sonar.[6]

7.2 SAS imagery

SAS data presented as grey-level imagery can be interpreted in much the same way as SSS imagery, but it should be noted that there are some significant differences. There are several effects caused by the fact that sonar returns are collected over a range of aspect angles, including [52]:

- specular reflections from strong scatterers (glint), more prominent than in SSS due to the increased range of sonar incidence angles, sometimes overwhelming non-specular returns from the seabed;
- differences in the appearance of complex features and resonant reflectors as viewed from different angles;
- shadows that move or change shape as the viewing angle varies; and
- aspect-dependence in bottom reverberations and multipath effects, particularly from sloping seabeds in shallow water areas.

Other differences in SAS imagery include:

- the larger inherent dynamic range of the data, in which highlights may be orders of magnitude more intense than in corresponding sidescan sonar images;
- wavenumber spectral data (including phase and amplitude information) allowing additional methods of processing to retrieve target structural information; and
- the sheer volume of data generated by a high-resolution system operating over wide swath widths.

Hansen *et al.* [52] have described strategies for processing SAS imagery containing these angular effects. Variation in the appearance of features with aspect angle often results in blurring of these features in images formed from the synthetic aperture. With appropriate processing, however, it is possible to mitigate these effects and even gain more information about targets, by studying these angular dependences. Hansen *et al.* used wavenumber processing to remove some glint effects from the imagery. Furthermore, they used the Fixed Focus Shadow Enhancement technique, based on a technique developed for SAR imagery

---

[6] Note that the stability constraints associated with SAS are considerably more stringent than those associated with high-resolution sidescan sonar, because signals sent and received at different times must be combined with the correct phase.

[53], to sharpen the shadows, which are modelled as moving targets. In this process, the shadow is made sharp, while the surrounding, non-moving imagery is blurred. This kind of processing has been demonstrated operationally [54]. It should be conducted as an intermediate step, after targets have been detected and prior to the classification of mine-like objects, to improve the classification performance.

Hagen and Hansen [55] have demonstrated that some of the difficulties associated with SAS processing can be overcome by effective design of the sonar hardware, in developments involving the Kongsberg Maritime HISAS 1030 sonar on the HUGIN 1000-MR AUV. Surface reverberation effects have been reduced by using a phased array transmitter, allowing the beam to be steered away from the sea surface. The addition of a second receiving array, parallel to the first and directly above it, has allowed estimation of the underwater topography via interferometric processing. The resulting topography is then employed to improve the focusing process, in comparison with what can be achieved by the assumption of a flat seabed. Their use of a relatively high frequency (for SAS) allows the recovery of shadows without undue loss of resolution. Finally, they and others have discovered that SAS processing reduces multipath effects, as different multipath signals arrive out of phase with direct arrivals and each other and are thereby integrated away during the processing.

Bell *et al.* [47] from the Heriot-Watt University group used their model-based approach, as described in Section 5.2 [37-40], to process SAS imagery. While the SAS imagery suffered from greater amounts of speckle and less distinct shadows than in corresponding SSS imagery, the cooperating statistical snakes algorithm was able to mark appropriate boundaries around the bright features and shadows, without being compromised by the speckle. In SSS imagery, target highlights are somewhat random in their appearance, and most of the information for classification of the targets is contained within the shadow. In the SAS imagery, however, Bell *et al.* found that target highlight areas of images also contain information useful for classifying the targets, due to the higher resolution of SAS.

The processing of SAS imagery is an area of ongoing research, to provide the best possible techniques for automatic detection and classification of significant seabed features.

8. Conclusion

Detection of significant objects such as mines has progressed significantly in recent years, with the emergence of reliable AUVs and the development of automated image processing techniques. The available techniques, while not mature, show great promise for reliable detections of mine-like objects in relatively uncluttered environments using a sidescan sonar or synthetic aperture sonar mounted on an AUV. With data from a high-resolution sidescan sonar, an object proud of the seabed can often be detected by the coincidence of an acoustic highlight and shadow in the image of the object, and the shape of the shadow indicates the geometry of the object. Synthetic aperture sonar has the advantage of allowing for high-resolution surveys out to a greater detection range. The shadows are generally less distinctive but the highlight resolution is often higher, so there is more emphasis on analysing the highlights, and not just the shadows, in classification of bottom objects detected using SAS systems.

Detection/classification routines can broadly be divided into two kinds: the operation of supervised and unsupervised algorithms. Supervised algorithms require training data to set up their operation; unsupervised algorithms do not. Supervised algorithms can be trained for the required detection task using images that are representative of the clutter backgrounds likely to be encountered, providing flexibility to cope with both straightforward and difficult detection tasks.

While supervised algorithms can be expected to perform better when there is a training data set appropriate for the test data (the data in which unknown mines must be detected), the task of obtaining an appropriate training data set is non-trivial. There must be mine-like objects in known locations, so that valid detections and false alarms can be identified, and the background clutter and reverberations should be typical and representative of those in the test data. There should also be enough training data so that anomalies in particular training images do not affect the overall detection performance. Training with data sets including atypical backgrounds and reverberations can actually impair the performance of a trained algorithm; it is better to restrict the training data to contain only backgrounds that are typical and representative of those encountered in the detection task at hand [17]. There is no clear measure of how appropriate the training data set is to the detection task at hand; human judgement may be required in making such decisions. When mine hunting in an area atypical of previously surveyed areas, significant time and resources may be required to gather training data, before the mine hunting begins in earnest. Once a suitable set of background imagery is obtained, the 'augmented reality' approach could be used [41] — artificially inserting mine shapes into background digital imagery, to alleviate any paucity of training data containing mine-like objects.

Unsupervised algorithms, set up with 'catch-all' detection processes, are simpler to implement, particularly as no training data set is required, but these algorithms cannot be expected to work as effectively in all circumstances as suitably trained algorithms.

The fusion of several different algorithms has been demonstrated to provide dramatic improvements in the probability of detection and correct classification of mine-like objects (for a given false alarm rate) over what can be achieved by any one of these algorithms individually.

It is recommended that DSTO continue investigating both unsupervised and supervised algorithms for detection of mine-like objects in sonar imagery from AUVs, to build up a set of trusted algorithms. This work will serve three main purposes:

1. enable the automatic processing of large volumes of data being acquired by DSTO and, during exercises, the RAN, so that features of interest can be easily discovered and interrogated;
2. enable comparative performance testing of different algorithms (or combinations of algorithms) as candidates for a post-processing aid for Defence operations; and
3. develop techniques for onboard processing on AUVs, to enable intelligent decision-making based on detected features.

Algorithms should be tested on a variety of data encompassing the range of environmental conditions likely to be encountered. It may be that different algorithms will perform better under different conditions of the sonar and the environment. Once several candidate algorithms are available, fusion of these algorithms is likely to improve the overall detection performance without increasing the false alarm rate. Testing of the best algorithms as decision aids can then take place, and comparisons with the performance of human analysts can be made.

The question remains as to whether automated detection and classification of mine-like objects will be trusted enough to be relied upon, without the need for a human operator to go back through all the data. How well will automated techniques work in areas of strong clutter or in rough seas causing strong surface reverberations? How well will they work when mines are partially buried? Questions such as these are difficult to answer, as they require extensive investigations. Overseas experience has suggested that it is very difficult to achieve a level of trust sufficient for automated systems to displace human analysts [2]. Mistakes made with the introduction of premature, poorly performing CAD/CAC systems are not easily forgotten. Even when the automatic detection/classification performance is better than for a human operator, such gains may not be recognised, as valid detections by the automated system tend to be regarded by human analysts as false alarms [2]. The introduction of CAD/CAC systems for post-processing of data must be very carefully managed to achieve the best possible outcomes.

In any event, the ability of AUVs to make onboard tactical decisions based on real-time processing of their imagery will greatly enhance their utility and performance in mine hunting operations. This is a role for automated systems that is not easily performed by human analysts, in the underwater domain in which high-bandwidth communications are difficult or impossible. CAD/CAC processing will also assist in the detection of changes over time in the distribution of mine-like objects, even in cluttered areas. Automated image processing will enable the detection of bottom objects to become more consistent and reliable and less labour-intensive than was previously possible.

9. Acknowledgement

The author wishes to thank Dr Stuart Anstee for his careful reading of the text and for the numerous suggestions to improve the content of this document.

10. References

1. S. Perry, *Applications of image processing to mine warfare sonar*, DSTO-GD-0237 (2000).
2. R.T. Kessel, *Apparent reliability – conditions for reliance on supervised automation*, Defence R&D Canada, DRDC Atlantic TM 2005-155 (2005).
3. R.T. Kessel, *On-screen alarms in computer-aided detection systems – combining signal detection, human factors and system design*, Defence R&D Canada, DREA TM 2001-184 (2001).
4. R.T. Kessel, *The burden of computer advice – using expert systems as decision aids*, Defence R&D Canada, DRDC Atlantic TR 2003-241 (2003).
5. V.L. Myers and M. A. Pinto, *Information theoretic bounds of ATR algorithm performance for sidescan sonar target classification*, Proc. SPIE **5807**, 130-140 (2005).
6. T. Hillier, *Digital side-scan: is this the end of TVG and AGC?*, Hydro International **12** (No. 2), 13-15 (2008).
7. P. Chapple, *Image processing for autonomous underwater vehicle operations*, presentation at Hydro 2007, Cairns, Australia, 21 – 24 November 2007.
8. P.B. Chapple, D.C. Bertilone, R.S. Caprari, G.N. Newsam, *Stochastic model-based processing for detection of small targets in non-Gaussian natural imagery*, IEEE Trans. Image Processing **10** (no. 4), 554-564 (2001).
9. S.G. Johnson, *Fast noise reduction for high-resolution sonar image enhancement*, Proc. IEEE OCEANS **1**, 331-336 (1992).
10. R.A. Carmona and S. Zhong, *Interior point methods for sea bottom image enhancement*, Proc. SPIE **3079**, 132-137 (1997).
11. W.G. Szymczak, W. Guo and J.C.W. Rogers, *Mine detection using variational methods for image enhancement and feature extraction*, Proc. SPIE **3392**, 286-296 (1998).
12. W.G. Szymczak and W. Guo, *Mine detection using model-trained multi-resolution neural networks and variational methods*, Proc. SPIE **3710**, 559-569 (1999).
13. Q. Huynh, N. Neretti, N. Intrator and G. Dobeck, *Image enhancement for pattern recognition*, SPIE **3392**, 306-314 (1998).
14. T. Aridgides, D. Antoni, M. Fernández and G. Dobeck, *Adaptive filter for mine detection and classification in side scan sonar imagery*, Proc. SPIE **2496**, 475-486 (1995).

15. T. Aridgides, M. Fernández and G. Dobeck, *Adaptive 3-dimensional range-crossrange-frequency filter processing string for sea mine classification in side-scan sonar imagery*, Proc. SPIE **3079**, 111-122 (1997).

16. R.T. Kessel, *Pass-fail performance testing for detection systems*, Defence R&D Canada, DREA TM 2001-205 (2002).

17. R. Manning, *Small object classification performance of high-resolution imaging sonars as a function of image resolution*, Proc. MTS/IEEE OCEANS '02, 2156-2163 (2002).

18. J.C. Hyland and G.J. Dobeck, *Sea mine detection and classification using side-looking sonar,* Proc. SPIE **2496**, 442-453 (1995).

19. G.J. Dobeck, J.C. Hyland and L. Smedley, *Automated detection/classification of sea mines in sonar imagery,* Proc. SPIE **3079**, 90-110 (1997).

20. G.J. Dobeck, *Fusing sonar images for mine detection and classification,* Proc. SPIE **3710**, 602-614 (1999).

21. M. Fernández, A. Aridgides and J. Bourdelais, *Algorithm for sonar-based signal identification*, Proc. IEEE OCEANS '93 Vol. **3**, 438-443 (1993).

22. T. Aridgides, M. Fernández and G. Dobeck, *Fusion of sea mine detection and classification processing strings for sonar imagery*, Proc. SPIE **4038**, 391-401 (2000).

23. T. Aridgides, M. Fernández and G. Dobeck, *Fusion of adaptive algorithms for the classification of sea mines using high resolution side scan sonar in very shallow water*, Proc. MTS/IEEE OCEANS **1**, 135-142 (2001).

24. S.G. Johnson and M.A. Deatt, *The application of automated recognition techniques to side-scan sonar imagery*, IEEE J. Oceanic Eng. **19** (No. 1), 138-144 (1994).

25. J. Huang, C. Ciany, M. Broadman and S. Doran, *Data fusion of computer aided detection/computer aided classification algorithms for classification of mines in very shallow water environments*, Proc. SPIE **4038**, 413-420 (2000).

26. C.M. Ciany and W. Zurawski, *Performance of computer aided detection/computer aided classification and data fusion algorithms for automated detection and classification of underwater mines*, Proc. CAD/CAC 2001, Halifax, Canada, Nov 2001.

27. M.R. Azimi-Sadjadi, D. Yao, Q. Huang and G.J. Dobeck, *Underwater target classification using wavelet packets and neural networks,* IEEE Trans. Neural Networks, **11** (No. 3), 784-794 (2000).

28. G.J. Dobeck, *Algorithm fusion for automated sea mine detection and classification*, Proc. MTS/IEEE OCEANS **1**, 130-134 (2001).

29. M.R. Azimi-Sadjadi, D. Yao, A.A. Jamshidi and G.J. Dobeck, *Underwater target classification in changing environments using an adaptive feature mapping,* IEEE Trans. Neural Networks, **13** (No. 5), 1099-1111 (2002).

30. T. Aridgides, M. Fernández and G.J. Dobeck, *Volterra fusion of processing strings for automated sea mine classification in shallow water*, Proc. SPIE **5794** (2005).

31. G.J. Dobeck, *A probabilistic model for score-based algorithm fusion*, Proc. MTS/IEEE OCEANS **3**, 2429-2434 (2005).

32. J.A. Fawcett, *Image-based classification of sidescan sonar detections*, Proc. CAD/CAC 2001, Halifax, Canada, Nov 2001.

33. J. Fawcett, M. Couillard, D. Hopkin, A. Crawford, V. Myers and B. Zerr, *Computer-aided detection and classification of sidescan sonar images from the CITADEL trial,* Intl Conf. On Detection & Classification of Underwater Targets, Proc. Institute Acoustics **29**, Part 6, 3-10 (2007).

34. M. Mignotte, C. Collet, P. Pérez and P. Bothemy, *Hybrid genetic optimization and statistical model-based approach for the classification of shadow shapes in sonar imagery*, IEEE Trans. Pattern Analysis & Machine Intelligence **22** (No. 2), 129-141 (2000).

35. M. Mignotte, C. Collet, P. Pérez and P. Bothemy, *Sonar image segmentation using an unsupervised hierarchical MRF model*, IEEE Trans. Image Processing **9**, (No. 7), 1216-1231 (2000).

36. R.A. Neill, *Optimizing display parameters for side-scan sonar operators*, Proc. 9th Commonwealth Def. Sci. Org Conference, Auckland, NZ 1991.

37. S. Reed, Y. Petillot and J. Bell, *An automatic approach to the detection and extraction of mine features in sidescan sonar*, IEEE J. Oceanic Eng. **28**, 90-105 (2003).

38. S. Reed, Y. Petillot and J. Bell, *Model-based approach to the detection and classification of mines in sidescan sonar*, Applied Optics **43** (No. 2), 237-246 (2004).

39. S. Reed, Y. Petillot and J. Bell, *Mine detection and classification in side scan sonar,* Sea Technology, Nov 04, 35-39 (2004).

40. S. Reed, Y. Petillot and J. Bell, *Automated approach to classification of mine-like features in sidescan sonar using highlight and shadow information*, IEE Proc. Radar, Sonar & Navigation **151** (No.1), 48-56 (2004).

41. E. Coiras, P.-Y. Mignotte, Y. Petillot, J. Bell and K. Lebart, *Supervised target detection and classification by training on augmented reality data*, IET Radar, Sonar & Navigation **1** (No. 1), 83-90 (2007).

42. E. Coiras, Y. Petillot and D.M. Lane, *Multiresolution 3-D reconstruction from side-scan sonar images*, IEEE Trans. Image Processing **16** (No. 2), 382-390 (2007).

43. P.-Y. Mignotte, E. Coiras, H. Rohou, Y. Pétillot, J. Bell and K. Lebart, *Adaptive fusion framework based on augmented reality training*, IET Radar, Sonar & Navigation **2**, 146-154 (2008).

44. R. Quintal, P. Dysart and R. Greene, *Automated Side-scan Data Analysis*, Hydro International **11** (No. 9), 22-25 (October 2007).

45. www.seebyte.com

46. R.J. Urick, *Principles of Underwater Sound*, McGraw-Hill (1983); *Sound Propagation in the Sea*, Peninsular Publishing (1982).

47. J.M. Bell, Y.R. Petillot, K. Lebart, P.Y. Mignotte and H. Rohou, *Target recognition in synthetic aperture and high resolution sidescan sonar*, Institution of Engineering & Technology Seminar on High Resolution Imaging and Target Classification, UK, 99-106 (2006).

48. www.km.kongsberg.com/ks/web/nokbg0240.nsf/AllWeb/90CBC5D82A8F2A6CC125721F003360C2?OpenDocument

49. www.ixsea.com/en/systems/002.002.001.001/shadows.html

50. www.qinetiq.com/home/newsroom/news_releases_homepage/2005/4th_quarter/saab_contract_signals.html

51. T.O. Sæbø, H.J. Callen and R.E. Hansen, *Bathymetric capabilities of the HISAS interferometric synthetic aperture sonar*, Proc. MTS/IEEE OCEANS 2007, 1-10 (2007).

52. R.E. Hansen, J. Groen and H.J. Callow, *Image enhancement in synthetic aperture sonar*, Intl Conf. On Detection & Classification of Underwater Targets, Proc. Institute Acoustics **29**, Part 6, 69-76 (2007).

53. T. Sparr, R.E. Hansen, H.J. Callow and J. Groen, *Enhancing target shadows in SAR images*, Electronics Letters **43** (No. 5), March 2007.

54. Ø. Midtgaard and P.E. Hagen, *Automatic target recognition for the HUGIN Mine Reconnaissance System,* Intl Conf. On Detection & Classification of Underwater Targets, Proc. Institute Acoustics **29**, Part 6, 29-36 (2007).

55. P.E. Hagen and R.E. Hansen, *Synthetic aperture sonar challenges*, Hydro International **12** (No. 4), 26-31 (2008).

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION DOCUMENT CONTROL DATA	1. PRIVACY MARKING/CAVEAT (OF DOCUMENT)

2. TITLE	3. SECURITY CLASSIFICATION (FOR UNCLASSIFIED REPORTS THAT ARE LIMITED RELEASE USE (L) NEXT TO DOCUMENT CLASSIFICATION)
Automated Detection and Classification in High-resolution Sonar Imagery for Autonomous Underwater Vehicle Operations	Document (U) Title (U) Abstract (U)
4. AUTHOR(S)	5. CORPORATE AUTHOR
Philip Chapple	**DSTO Defence Science and Technology Organisation** **PO Box 1500** **Edinburgh South Australia 5111 Australia**

6a. DSTO NUMBER	6b. AR NUMBER	6c. TYPE OF REPORT	7. DOCUMENT DATE
DSTO-GD-0537	AR-014-199	General Document	December 2008

8. FILE NUMBER	9. TASK NUMBER	10. TASK SPONSOR	11. NO. OF PAGES	12. NO. OF REFERENCES
U-490-6-328-1	NAV 07/073	CANHMG	31	55

13. URL on the World Wide Web	14. RELEASE AUTHORITY
http://www.dsto.defence.gov.au/corporate/reports/DSTO-GD-0537.pdf	Chief, Maritime Operations Division

15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT

Approved for public release

OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE, PO BOX 1500, EDINBURGH, SA 5111

16. DELIBERATE ANNOUNCEMENT

No Limitations

17. CITATION IN OTHER DOCUMENTS Yes

18. DSTO RESEARCH LIBRARY THESAURUS http://web-vic.dsto.defence.gov.au/workareas/library/resources/dsto_thesaurus.shtml

Sonar, Detection, Target classification, Image analysis

19. ABSTRACT
Autonomous Underwater Vehicles (AUVs) are increasingly being used by military forces to acquire high-resolution sonar imagery, in order to detect mines and other objects of interest on the seabed. Automatic detection and classification techniques are being developed for several reasons: to provide reliable and consistent detection of objects on the seabed; to free human analysts from time-consuming and tedious detection tasks; and to enable autonomous in-field decision-making based on observations of mines and other objects. This document reviews progress in the development of automated detection and classification techniques for side-looking sonars mounted on AUVs. Whilst the techniques have not yet reached maturity, considerable progress has been made in both unsupervised and supervised (trained) algorithms for feature detection and classification. In some cases, the performance and reliability of automated detection systems exceed those of human operators.